Analysis of electric-field-induced spin splitting in wide modulation-doped quantum wells

U. Ekenberg*
School of Information and Communication Technology, Royal Institute of Technology, SE-16440 Kista, Sweden

D. M. Gvozdić
Faculty of Electrical Engineering, University of Belgrade, P.O. Box 35-54, 11120 Belgrade, Serbia
(Received 29 December 2007; revised manuscript received 17 October 2008; published 18 November 2008)

We analyze the proper inclusion of electric-field-induced spin splittings in the framework of the envelope function approximation. We argue that the Rashba effect should be included in the form of a macroscopic potential as diagonal terms in a multiband approach rather than the commonly used Rashba term dependent on k and electric field. It is pointed out that the expectation value of the electric field in a subband is sometimes not unique because the expectation values can even have opposite signs for the spin-split subband components. Symmetric quantum wells with Dresselhaus terms and the influence of the interfaces on the spin splitting are also discussed. We apply a well established multiband approach to wide modulation-doped InGaSb quantum wells with strong built-in electric fields in the interface regions. We demonstrate an efficient mechanism for switching on and off the Rashba splitting with an electric field being an order of magnitude smaller than the local built-in field that determines the Rashba splitting. The implications of our findings for spintronic devices, in particular the Datta-Das spin transistor and proposed modifications of it, are discussed.

DOI: 10.1103/PhysRevB.78.205317 PACS number(s): 71.70.Ej, 73.21.Fg, 85.75.Hh

I. INTRODUCTION

The interaction of a spin with a magnetic field or magnetic ions has been studied extensively over the years. A surprisingly efficient mechanism to produce a spin splitting is utilizing the combination of inversion asymmetry and spin-orbit interaction.[1] One convenient way to control the spin splitting is the Rashba effect which results from structure inversion asymmetry (SIA).[2] An applied or built-in macroscopic electric field is seen in the frame of a moving electron as having a magnetic-field component and yields a spin splitting. In this way one can utilize many of the mechanisms from conventional electronics which is mainly controlled by electric fields. The idea of using the spin of the carriers in addition to its charge has resulted in a research area called spintronics.[3] Another spin splitting mechanism resulting from the lack of inversion symmetry of the zinc-blende lattice [bulk inversion asymmetry (BIA)] is called the Dresselhaus effect.[4] Both the Rashba and the Dresselhaus effects are frequently included via terms linear in the wave vector k. However, they are the lowest-order terms of more accurate expressions that are obtained from multiband envelope function theory.[1]

In Sec. II we will recapitulate the foundations of the commonly used envelope function approximation[5,6] in order to set the ground for an analysis of the proper inclusion of SIA and BIA within the framework of this approximation. We make here the important distinction between slowly and rapidly varying potentials. In Sec. III we apply the multiband theory to an interesting system, wide n-type modulation-doped (MD) quantum wells (QWs) with strongly nonuniform electric fields. Here there are strong built-in electric fields of opposite signs at the two interfaces and one can expect a strong Rashba effect. It has frequently been assumed that one needs a strong applied bias to get a substantial Rashba splitting but interesting things can happen also for small or moderate bias. We show how the built-in electric fields in modulation-doped quantum wells can be utilized while applying a much smaller external field. For very small asymmetry interesting anticrossing phenomena occur. Furthermore, the spin splitting due to the Dresselhaus effect in symmetric quantum wells is found to be qualitatively different in modulation-doped quantum wells compared to square wells. In Sec. IV we discuss the implications of our results for spintronic devices, in particular the Datta-Das spin field effect transistor.[7] It has for a long time been considered as a prototype of a spintronic device but unfortunately the efforts to implement it in practice have not yet been very successful. Finally, in Sec. V we discuss the results and conclude.

II. THEORY

The envelope function approximation (EFA) has been widely used during several decades. Under the name effective-mass theory, it was first applied to shallow-impurity states in bulk semiconductors.[8,9] The starting point is that the problem with the band structure in the pure bulk material is assumed to have been solved. According to Bloch's theorem the total wave function for band n is given by

$$\psi_{n\mathbf{k}}(\mathbf{r}) = e^{i\mathbf{k}\cdot\mathbf{r}} u_{n\mathbf{k}}(\mathbf{r}), \tag{1}$$

where $u_{n\mathbf{k}}(\mathbf{r})$ has the periodicity of the lattice. We then introduce a perturbation $U(\mathbf{r})$. An essential assumption in the derivation is that it should be slowly varying on the scale of the lattice constant. This assumption does not always hold in the cases where the EFA has been applied. The advantage of the EFA is that one can avoid the explicit inclusion of the cell periodic potential. Only the slowly varying perturbation $U(\mathbf{r})$ enters a Schrödinger-like equation. With the perturbation the total wave function can be expanded

$$\psi(\mathbf{r}) = \sum_n f_n(\mathbf{r}) u_{n\mathbf{k}}(\mathbf{r}), \tag{2}$$

where the summation is over all the energy bands. $f_n(\mathbf{r})$ is called an envelope function and the EFA gives a simple prescription for the effective Hamiltonian operating on the envelope function f_n only,[5,6]

$$H = E(-i\nabla) + U(\mathbf{r}). \tag{3}$$

The kinetic-energy operator E is obtained from the bulk band structure $E(\mathbf{k})$. In general it is a matrix whose eigenvalues gives the energy-band dispersion in the bulk. For the perturbed problem $\mathbf{k}$ is replaced by $-i\nabla$, where ∇ is the gradient operator. In the case of a quantum well grown along the z direction, it is sufficient to replace k_z by the operator $-i\partial/\partial z$ while k_x and k_y remain good quantum numbers. The perturbation potential $U(\mathbf{r})$ is added along the diagonal of the matrix.

For the impurity case the potential is not slowly varying in the unit cell containing the impurity. For s states having a finite amplitude at the origin, this sometimes gives important deviations (often called central-cell corrections) from the predictions of the EFA in its simplest form.[10] The ground state for a donor in Si is, for example, split up into three levels where the energy separation between these levels can be comparable to the predicted ground-state energy. For p and d states in Si the EFA works excellently and it also works well for the ground state in direct-gap semiconductors.[11]

The EFA has also been used frequently for quantum well heterostructures. Here the potential changes rapidly near the interface between two materials. The range of this potential change is of the order of the lattice constant. Thus, at first sight, it appears that the EFA would not be applicable. In spite of this, the agreement between its predictions and experimental results has turned out to be quite good if the EFA is applied properly without unnecessary approximations. The reason for this was examined by Burt[12] in a series of papers and led to a new set of boundary conditions[13,14] nowadays called the Burt-Foreman boundary conditions. These boundary conditions were first derived for the 6×6 Hamiltonian describing the valence bands[13] and later a prescription for extending it to the 8×8 Hamiltonian including the conduction band was given.[14] Previously it was common to symmetrize the order between operators and spatially varying parameters but the Burt-Foreman boundary conditions were derived with consideration of the cell-periodic wave functions $u_{n\mathbf{k}}(\mathbf{r})$. Meney *et al.*[15] analyzed different envelope function approaches and found support for the Burt-Foreman boundary conditions. Burt's analysis also explained why the EFA works quite well even for narrow quantum wells. However, it should be kept in mind that caution is necessary when applying the EFA to interface regions. In the EFA the interfaces have usually been taken as abrupt steps. A rapidly but continuously varying potential has also been considered by Stern and Das Sarma[16] but the influence on the subband energies was found to be quite small.

The summation in Eq. (2) should in principle be over all the bands. In practice one selects a finite number of important bands whose interaction is included exactly in the matrix while the "remote" bands are included perturbatively.[17] A larger number of bands included in the matrix gives an accurate description in a larger k range. A common choice that we apply in this paper is to include the conduction, heavy-hole, light-hole, and split-off bands in an 8×8 matrix. For a symmetric structure this includes a twofold spin degeneracy.

For even more accurate results, 14×14 and 16×16 matrices have been considered.[18] Wissinger *et al.*[19] have performed calculations for an asymmetric GaAs quantum well both with the 8×8 and the 14×14 matrices, and compared to Raman-scattering experiments.[20] The deviation between the two models was rather small and comparable to the deviation from the experimental results. In the present case we consider InGaSb with a small band gap where the 8×8 matrix is expected to be a good approximation. On the other hand, it can be sufficient to use 6×6 matrices in which the conduction band or the split-off band is among the remote bands. The frequently applied Luttinger-Kohn Hamiltonian[5,6] includes the heavy-hole and light-hole bands in a 4×4 matrix. For a description of electron subbands, it is convenient to use a one-band (two-component) approximation in which all the other bands are considered as remote and included via a modification of the free-electron mass to an effective mass. As we will see below, inclusion of spin effects in a one-band model leads to some complications.

The recapitulation above of the essence of the EFA serves as the basis for analyzing the inclusion of asymmetry-induced spin phenomena. The inclusion of the Dresselhaus effect[4] seems clear. It stems from the microscopic structure of the bulk material, influences the cell-periodic part $u_{n\mathbf{k}}$ of the wave function, and results in a modified bulk band structure. Thus it is appropriate to include it as k-dependent terms in the matrix which, after the replacement $\mathbf{k}\rightarrow-i\nabla$, becomes the kinetic-energy operator. Several terms of different order in k enter the 8×8 Hamiltonian.[1] For electron subbands the lowest-order term is linear in k:

$$H_D = \beta(k_x\sigma_x - k_y\sigma_y), \tag{4}$$

where σ_x and σ_y are Pauli matrices, and β is a material constant giving the strength of the Dresselhaus effect.

The inclusion of the Rashba effect[2] is less uncontroversial. It stems from a slowly varying macroscopic electric field and, according to the principles of the EFA, it should be included as a z-dependent potential along the diagonal of a matrix of sufficient size. The Rashba effect in p-channel Si metal-oxide-semiconductor field-effect transistor (MOSFET) structures was already included in this way in the 1970s and good agreement with experiment was found.[21] Using a multiband matrix for the kinetic energy, the inclusion of an asymmetric potential results in a spin splitting for finite values of the in-plane wave vector without inclusion of any special k-dependent terms [cf. Eq. (5) below]. The spin-orbit interaction is implicitly included via the coefficients of the k-dependent elements in the matrix. They contain matrix elements of the spin-orbit interaction with respect to the cell-periodic wave functions $u_{n\mathbf{k}}$ and can be evaluated theoretically.[4] However, a higher accuracy can often be obtained from cyclotron resonance experiments,[22] and in prac-

tice experimentally determined effective masses and Luttinger parameters[6] are inserted if they are available.

Cyclotron resonance experiments for a two-dimensional hole gas at a modulation-doped GaAs/AlGaAs interface were performed by Störmer *et al.*[23] The roughly triangular potential with a strong Rashba effect gave rise to two clearly different masses ascribed to the two components of the spin-split heavy-hole subband. These results were in very good agreement with calculations with the Luttinger-Kohn Hamiltonian, where the measurable energies of the allowed Landau-level transitions were explicitly calculated.[24]

An alternative to the numerical solution of a multiband problem with suitable boundary conditions is a process called "downfolding." Starting from a multiband matrix, one can derive various 2×2 Hamiltonians.[25–27] The commonly used Rashba term can be derived as the lowest-order term.[1] For the electric field in the z direction, it becomes[2]

$$H_R = \alpha(k_x\sigma_y - k_y\sigma_x). \tag{5}$$

Here α is often used as an input parameter taken from experiment. Several experiments (see, e.g., Refs. 20, 28, and 29) have aimed at determining this Rashba coefficient for different materials. Various theoretical expressions have been derived for α. A simple expression that we will use as a reference is[25]

$$\alpha = \frac{\hbar^2 e\Delta(2E_g+\Delta)}{2mE_g(E_g+\Delta)(3E_g+2\Delta)}\langle\varepsilon\rangle, \tag{6}$$

where $\langle\varepsilon\rangle$ is the expectation value of electric field in the quantum well and the barriers.

The Rashba term is not really consistent with the principles of the EFA. It is a kind of hybrid including both the wave vector k and the potential. The problem arises from the fact that the s-like conduction band gets its spin-orbit coupling from the interaction with the valence bands, which are included in this approximation among the "remote" bands. A further problem is that Eq. (6) implicitly assumes that the electron subband has a well-defined expectation value but as will be discussed below, the two components of the spin-split subband (henceforth denoted spin subbands) can have clearly different expectation values.

A special problem is how the interfaces of a quantum well should be included. In an asymmetric quantum well the penetration of the wave function into the left and right barriers becomes different. At a first glance it seems natural to treat the complete conduction-band profile as the relevant potential. Zawadzki and Pfeffer[26] have included the conduction-band offsets in $\langle\varepsilon\rangle$ [cf. Eq. (6)] and denoted it the "average electric field." Using the fact that no force acts on a bound state, it has been argued[1,26] that the contribution from the interfaces would largely cancel that of the electric field in the quantum well. However, in this respect it is important to distinguish between the total wave function and the envelope functions for which this argument does not necessarily hold. The different effective masses in well and barrier, and spin dependent boundary conditions make this average electric field yield a nonzero but small contribution to the Rashba spin splitting.

Lassnig[27] argued that the valence-band profile, including the band offsets, determines the Rashba effect for conduction electrons. This gives an interface contribution of the same sign as that of the electric field in the quantum well and barriers. This interface contribution has been evaluated analytically by downfolding the other bands on the conduction band and resulted in matrix elements of the steplike valence-band edge.[26] However, this approach of treating the interfaces like infinite electric fields seems dubious against the discussion above, where it was pointed out that the interfaces are a weak point in the EFA. A crucial nontrivial factor in the downfolding procedure is how the order between differential operators and spatially varying material parameters should be chosen to be compatible with the Burt-Foreman boundary conditions.[13,14] So far downfolding approaches have usually used the *ad hoc* operator symmetrization that many workers have abandoned in multiband approaches. We will analyze the interface contribution below in our multiband approach using Eqs. (5) and (6).

This interface contribution is fundamentally different from another interface contribution where the microscopic structure at the interface is taken into account as an additional source of inversion asymmetry. The proper inclusion of this short-range potential in the framework of envelope function theory is not trivial and has been subject to debate.[30] The effect is particularly strong in "no-common-atom interfaces" such as InAs/GaSb where the two constituents have no atom in common.[31] We have neglected it in the present calculations.

Equations (5) and (6) predict a linear increase in the Rashba spin splitting with the in-plane wave vector k. Numerical results[25,32] give a Rashba splitting that is nonlinear and can even be nonmonotonic. Yang and Chang[32] have recently published numerical and approximate analytical solutions to the Rashba effect. By inserting the actual subband energy $\varepsilon(k)$ instead of the bulk band-edge energies as in simpler models, their analytical model qualitatively reproduces the nonlinear k dependence. However, the analytical model overestimates the spin splitting. Their numerical results for an $In_{0.53}Ga_{0.47}As$ quantum well between $In_{0.52}Al_{0.48}As$ barriers agree very well with what we obtain in our approach. The quantitative agreement for $Hg_{0.74}Cd_{0.26}Te/HgTe$ structures is less good but the qualitative behavior of an increase followed by a decrease in the spin splitting is reproduced. We find a similar nonmonotonic behavior for an InGaSb/InAlSb structure with the same composition as in this paper but with a uniform electric field. On the other hand, for a GaAs/AlGaAs quantum well we find a monotonic but clearly nonlinear behavior, in agreement with Ref. 25. This is consistent with the explanation[32] that the nonmonotonic k dependence is caused by the increased energy separation to the light-hole and split-off bands, an effect that is less pronounced in GaAs with a fairly large band gap. It can be noted that the dependence of ΔE on k is similar to the dependence for holes in Ref. 33 of Δk at the optimal energy on the electric field which was explained in a similar way.

We have started from the properties of the bulk materials and, via the boundary conditions, obtained the subband dispersions. An alternative approach is to solve the subband

problem for $k=0$ and expand the wave functions for finite k in eigenfunctions for $k=0$ (the mini-$\mathbf{k}\cdot\mathbf{p}$ method).[34] This approach can be illuminating also for spin phenomena.[1] For a modulation-doped interface it was found that, however, the convergence with respect to number of basis functions was surprisingly slow.[35]

III. RESULTS

We here consider an interesting system for studying spin effects, an n-type wide modulation-doped quantum well. Due to the attraction to the ionized donors in the barriers, we obtain two weakly interacting electron gases mainly localized to the interface regions. The Rashba effect has often been studied experimentally and theoretically in structures where the electric field is uniform or close to uniform. In wide modulation-doped quantum wells, the electric field is strongly nonuniform [see Fig. 1(a)]. In each interface region there is a strong electric field which, according to Poisson's law, is proportional to the charge transferred from the donors in the barrier to the quantum well. This field is capable of producing a substantial Rashba effect.

The modulation-doped quantum well is a very versatile system. The degree of interaction between the electron gases can be controlled by the well width and the carrier concentration. The asymmetry can be regulated by an applied bias or by choosing unequal spacer layer widths.

In this paper we consider an 80 nm wide n-type $In_{0.74}Ga_{0.26}Sb$ quantum well with $In_{0.7}Al_{0.3}Sb$ barriers. The growth direction is [001]. The input parameters are given in Table I. Among the common III-V semiconductors InSb has the strongest spin-orbit coupling. By mixing in Ga and Al, respectively, we obtain a quantum well with an almost equally strong spin-orbit coupling. This InGaSb/InAlSb system has recently been studied experimentally by Akabori *et al.*[36] Similar effects should occur also for smaller well widths but the effects we want to display become quite clear for a well width of 80 nm. We here make a much more thorough analysis than in a recent preliminary paper.[37]

We use the well established approach with an 8×8 matrix approach with a minor modification (see Table Caption I) of the matrix given as Table C.8 in Ref. 1. We apply the Burt-Foreman boundary conditions,[12–14] and to avoid spurious solutions, we use a quadrature method in which unphysically large k values do not enter.[38] The subband problem is solved self-consistently in the Hartree approximation.

For reference we first consider a symmetric quantum well. If we, for the moment, ignore the Dresselhaus effect, we have a ground state with a symmetric wave function and a small energy separation to an excited state with an antisymmetric wave function. If the modulation-doped quantum well is wide enough that the electron gases can be considered as noninteracting, we seem to have a paradox. Each electron gas is in a strongly asymmetric potential and a strong Rashba effect can be expected. On the other hand, if the quantum well is considered as a whole, the potential is symmetric and a twofold spin degeneracy should result [see Fig. 1(a)]. This was sorted out previously[39] for a p-type quantum well, where the Rashba effect can be made several orders of magnitude larger than for electrons (cf. Ref. 33). However, the qualitative features are the same for the n-type quantum well. We recapitulate here the essentials and refer to Ref. 39 for details.

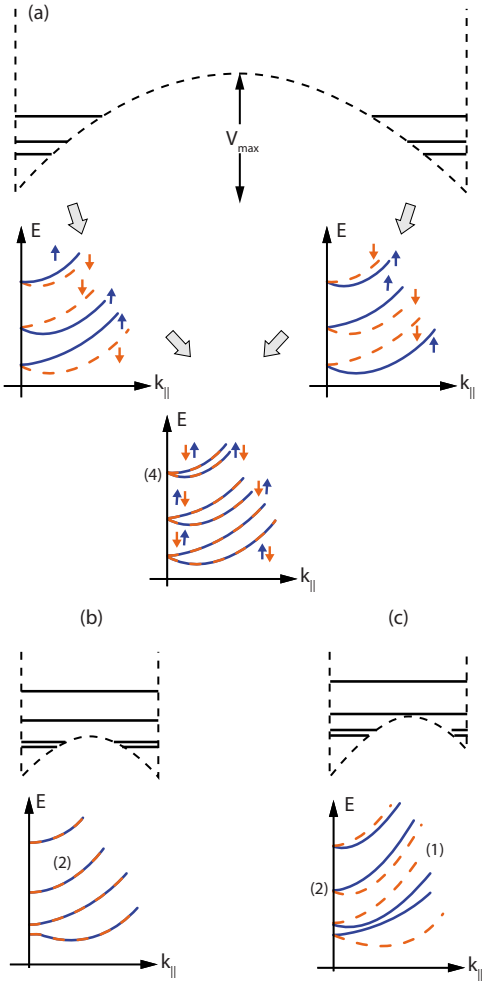

FIG. 1. (Color online) Schematic subband dispersions and conduction-band profiles for n-type MD quantum wells. The numbers in parentheses denote the degeneracy of the subbands for $k\neq 0$ and $k=0$. In these figures the Dresselhaus splitting has been ignored. (a) Here the well is wide enough that the two electron gases at the interfaces can be considered as noninteracting. To the left and the right we display the subband dispersions when the interface regions are considered separately. The middle figure shows the case when the quantum well is considered as a whole. (b) Symmetric MD quantum well of intermediate width. The previous fourfold degeneracy at $k=0$ partially lifted due to the interaction between the electron gases. (c) Asymmetric MD quantum well. Due to the overall asymmetry the degeneracy is lifted for $k\neq 0$.

First it should be noted that the signs of the electric fields at the interfaces are opposite to each other. If we label the upper spin subband at the left interface by spin up, the corresponding spin subband at the right interface should be labeled spin down. Looking now at the whole quantum well with two electron gases, the lower spin subband has equal amounts of spin up and spin down, and the twofold spin degeneracy expected for a symmetric potential prevails [lower subband structure in Fig. 1(a)]. This implies that what

TABLE I. Input parameters used in the present work for an $In_{0.74}Ga_{0.26}Sb$ quantum well surrounded by $In_{0.7}Al_{0.3}Sb$ barriers. We give the "true" Luttinger parameters although modified Luttinger parameters are included in the 8×8 matrix.[1] The conduction-band offset ΔE_c and valence-band offset ΔE_v between the two materials are also given. Mass parameters are given in units of the free-electron mass. Energies are given in electron volts. The coefficients C and B_{8v}^{+} describing the Dresselhaus effect are given in eV Å and eV Å^2, respectively. We use the reasonable approximations $B_{7v} \approx B_{8v}^{+}$ and $B_{8v}^{-} \approx 0$ that are justified by Table 6.3 in Ref. 1. Finally we give the static dielectric constant ϵ.

Parameter	InGaSb		InAlSb
m_e	0.015 95		0.042 21
γ_1	23.87		11.652
γ_2	10.01		4.110
γ_3	11.34		5.388
E_p	24.27		21.92
E_g	0.3068		0.790
Δ	0.777		0.717
C	−0.006 715		−0.006 524
B_{8v}^{+}	20.596		7.21
ΔE_c		0.3344	
ΔE_v		0.1487	
ϵ	17.17		16.00

looks like a spin-split subband in the single interface case should actually be considered here as two separate subbands. At $k=0$ we have a fourfold degeneracy because we have two spin directions and two electron gases. If we now consider a narrower but still symmetric quantum well such that the electron gases start to interact, the main and somewhat unexpected effect is that the degeneracy at $k=0$ is partially lifted, and for all k values, we have two closely spaced subbands, each with a twofold spin degeneracy [Fig. 1(b)].

This effect can possibly be related to the splitting recently obtained by Bernardes *et al.*,[40] who derived some off-diagonal matrix elements between wave functions of even and odd parity. This gives rise to a splitting qualitatively similar to the splitting described above when the electron (or hole) gases start to interact. However, the mechanism for the splitting derived in Ref. 40 is the spin-orbit interaction. Since well widths and other input parameters are not given in this paper, we cannot investigate if this effect is implicitly included in our multiband approach.

The twofold spin degeneracy becomes lifted for finite k if the potential of the quantum well is made asymmetric [Fig. 1(c)]. This is the case treated in this paper. The behavior described above is confirmed by numerical calculations.[39]

We next include the Dresselhaus terms in the matrix but keep the quantum well potential symmetric. The Dresselhaus effect becomes rather different in a wide modulation-doped quantum well compared to a square well.[1] In Fig. 2(a) we display the energy spin splitting as a function of wave vector. It is seen that it first rapidly increases but then decreases, and for larger wave vectors (approximately above the Fermi wave vector k_F), it stays rather constant at a low value. Our

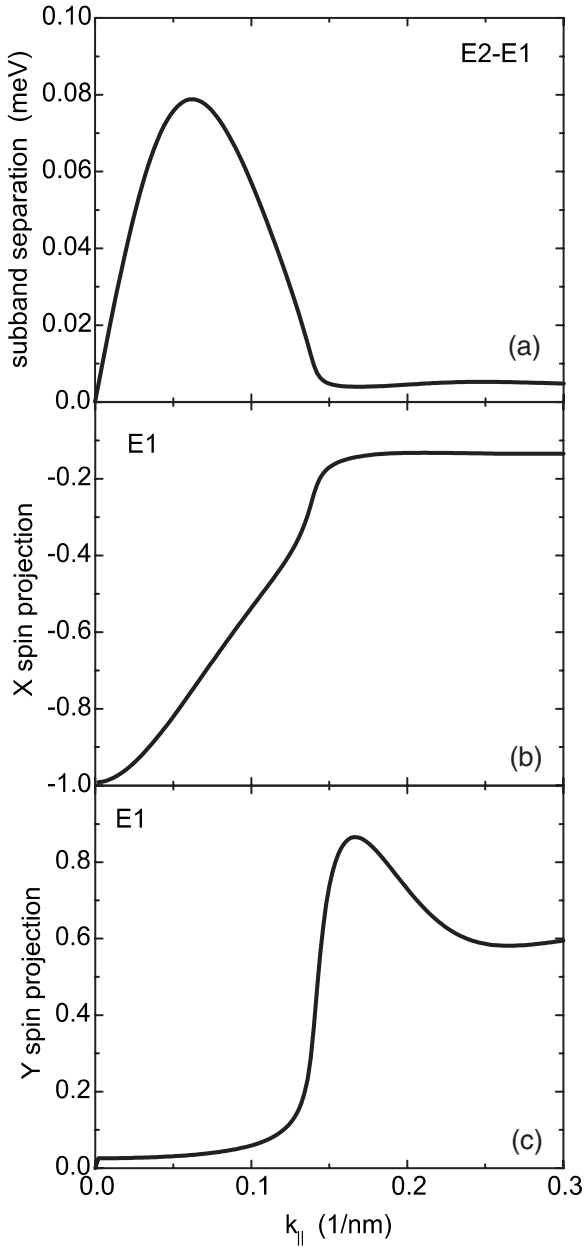

FIG. 2. (a) Energy subband splitting for an 80 nm n-type modulation-doped symmetric InGaSb quantum well. The wave vector is in the [10] direction in the two-dimensional Brillouin zone. [(b) and (c)] k dependence of the x and y components, respectively, of the spin vector. Here both the spacer layer widths are 45 nm and the electron density is 6.8×10^{11} cm^{-2}.

results imply that the change in spin splitting between a symmetric and an asymmetric wide modulation-doped quantum well normally is dominated by the Rashba effect. In Figs. 2(b) and 2(c) we show the k dependence of the x and y components of the expectation value of the spin vector.[1] The absolute value of the x component decreases rapidly and becomes small above k_F. This is in contrast to a square well where it stays constant.[41] The y component has the reverse behavior: small for small k values and increases rapidly near k_F. Thus the spin direction changes from the x direction to the y direction as k increases along the [10] direction. For

TABLE II. Spin splittings for different electric fields in various approximations. ε_{ave} is the voltage across the quantum well divided by the well width and ΔE_{ave} is the energy spin splitting obtained by inserting this electric field into Eq. (6). $\langle \Delta E \rangle_i^{excl}$ are the results when the expectation value of the electrostatic field in the layers excluding interface contributions (see text) averaged over filled states for spin subband i is inserted into Eq. (6). $\langle \Delta E \rangle_i$ are the corresponding results with inclusion of interface contributions from the valence-band offsets [Eq. (7)]. ΔE_{num} is our numerical result with the spin splittings evaluated at the Fermi wave vector $k_F \approx 0.13$ nm^{-1}. The last row corresponds to the situation in Fig. 3.

ε_{ave} (kV/cm)	ΔE_{ave} (meV)	$\langle \Delta E \rangle_{1\downarrow}^{excl}$ (meV)	$\langle \Delta E \rangle_{1\uparrow}^{excl}$ (meV)	$\langle \Delta E \rangle_{1\downarrow}$ (meV)	$\langle \Delta E \rangle_{1\uparrow}$ (meV)	ΔE_{num} (meV)
0.297	0.012	0.862	−0.292	1.384	−0.495	0.720
1.242	0.052	1.112	0.968	1.768	1.491	2.158
3.012	0.135	1.357	1.328	2.145	2.040	2.498
4.137	0.190	1.485	1.459	2.336	2.231	2.646

sufficiently large k we find a localization of the wave functions to one of the interface regions, similarly to what is shown below for the Rashba effect.

In a quantum well with small asymmetry, the Rashba splitting can be comparable to the splitting between the symmetric and the antisymmetric states. In addition to a gradual transfer of wave-function amplitude to one of the interface regions, there will be interesting anticrossing phenomena, especially when k is in the [11] direction.[42,43] In Table II we show the spin splittings for some values of the bias in various cases. Especially for small bias the expectation values can be quite different and even have opposite signs for the two spin subbands. This presents a fundamental problem in applying Eqs. (5) and (6) for more complex situations. Many downfolding approaches lead to expectation values of the electric field with respect to a subband rather than a spin subband. If one wants to treat this effect perturbatively, it would be appropriate to apply degenerate perturbation theory. However, since we find that a small perturbation can give rise to a substantial effect on the wave functions, the use of perturbation theory appears dubious in the present case. The anticrossing phenomena including the contrasting results for k in the [11] direction have been examined more closely elsewhere.[42,43]

In this paper we focus on the case with a fairly small but sufficiently large asymmetry that each wave function becomes almost completely localized to one of the interface regions. In the present case a bias over the quantum well [henceforth denoted quantum well bias (QWB)] of 33 mV is sufficient to reach this situation. It corresponds to a rather small average electric field of 4.1 kV/cm (last row in Table II). It gives an energy separation of 9.2 meV at $k=0$ between the lowest and next lowest subband pairs. It is shown in Fig. 3(b) that it gives a spin splitting that is an order of magnitude larger than for the same uniform electric field. By comparing columns 2 and 7 in Table II, we can observe for different biases how the spin splitting is enhanced in modulation-doped quantum wells compared to undoped quantum wells with the same QWB. For the largest (last row) and the smallest biases (first row), we have an enhancement by a factor of 14 and 60, respectively.

We thus have a modified and very efficient mechanism to apply a moderate QWB, and take advantage of the much stronger built-in electric field to obtain a substantial Rashba splitting. A qualitative explanation of the enhancement can be seen in Fig. 3(a) where the ground-state wave function has become localized to one of the interface regions. There the electric field becomes quite strong and it is this local field that determines the size of the spin splitting.

We note in Table II than we can reach a spin splitting of 2.6 meV at a wave vector of 0.13 nm^{-1}. According to Eq. (5) this corresponds to value as large as 200 meV Å for the Rashba parameter α. This compares well in comparison to experimentally determined α values.[20,28,29] In addition, it is essential that in our case this α value is reached with a moderate bias.

We will now investigate if we can reproduce this strong enhancement using the common Rashba model together with Eq. (6). As a first step we insert into Eq. (6) the expectation value of the electric field in the well and barriers ignoring any interface contributions (columns 3 and 4). It can be expected to be enhanced by the localization of the wave function. In the last row of Table II the spin subbands have almost the same expectation value. However, this procedure gives a clearly smaller Rashba splitting than obtained in our numerical calculations. We have found that inclusion of BIA has a small effect on the results in Table II.

As a next step it is natural to examine if the discrepancy can be explained by interface contributions. After solving the multiband problem in our approach, it is straightforward to evaluate the expectation value of the electric field in the layers and that of the steps at the interfaces separately. If the Rashba model Eqs. (5) and (6) is to be used, the most reasonable approach seems to be to follow Ref. 27 and take the expectation value of the valence-band profile. The interface contributions involve the derivative of the discontinuities in the valence bands and yield Dirac delta functions in the integrals that nevertheless can be evaluated. The result becomes

$$\langle \varepsilon \rangle_{\text{interfaces}} = \Delta E_v (|\psi(-a)|^2 - |\psi(a)|^2)/e, \tag{7}$$

where the interfaces are taken at $z = \pm a$. It is seen in Table II that the expectation values are increased by about 50% when the contribution from the interface steps is added. This is compatible with the analytical results by Yang and Chang[32]

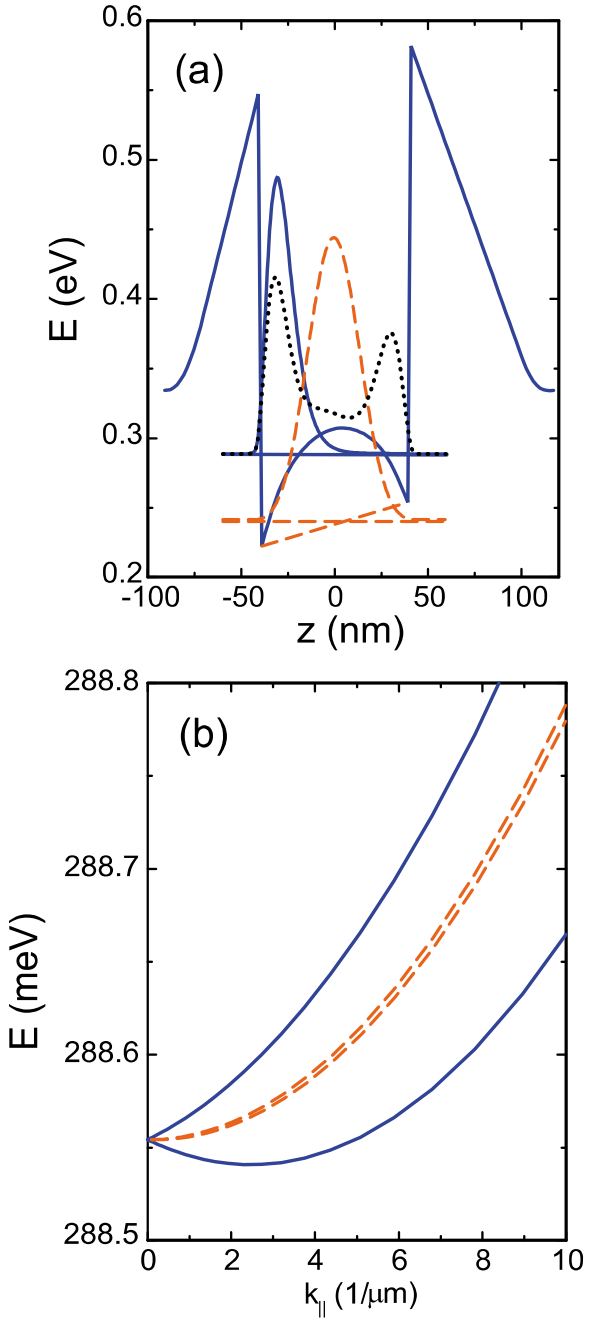

FIG. 3. (Color online) (a) Potential, squared wave function and charge density, and (b) subband dispersion along the [10] direction for the lowest subband pair in an 80 nm InGaSb quantum well. The quantum well bias (potential difference between the interfaces) is 33 mV. Dashed lines: uniform electric field; solid lines: modulation-doped quantum well with an electron density of 6.8×10^{11} cm^{-2}. For the latter case the dotted line shows the charge distribution.

who concluded that the interface contribution to the Rashba splitting was typically a factor of two smaller than that of the electric field in the quantum well.

We have also calculated the expectation value of the conduction-band profile to obtain the "average electric field."[26] The contribution from the conduction-band offsets is of opposite sign compared to that from the electric field in the quantum well and results in a net contribution of 10%–15% of our numerical value. In Ref. 26 a contribution of only 3% was obtained.

To examine the interface effect we have also replaced the barrier material by $In_{1-x}Al_xSb$ of different compositions. This changes the conduction- and valence-band offsets, the effective electron mass, and the nonparabolicity. If the interface contribution were very sensitive to these parameters, this should give a clear effect. However, in our calculations we only find a minor change. For example, if we change the barrier material from $In_{0.7}Al_{0.3}Sb$ to $In_{0.5}Al_{0.5}Sb$ (ignoring introduced strain), we obtain a change of the spin splitting by about 1%.

The failure to reproduce our numerical results using Eqs. (5) and (6) and various expectation values raises the question if more elaborate downfolding approaches can yield better agreement or if approaches using expectation values simply are insufficient. A more comprehensive comparison between various multiband and downfolding approaches will be published elsewhere. It is conceivable that two potentials having the same expectation value of the electric field can yield different Rashba spin splittings and that the spatial variation in the electric field must be taken into account in a multiband approach rather than basing the calculations on some kind of expectation values.

IV. IMPLICATIONS FOR SPINTRONIC DEVICES

The strong enhancement of the Rashba splitting described in Fig. 3 due to modulation doping can be expected to have important implications for several spintronic devices based on the Rashba effect. For the moment we focus on one of the best known spintronic devices, the spin transistor proposed by Datta and Das,[7] including proposed modifications of it. We will return to the problems encountered to make it function and first address the question: If it can be made to function, does it have the potential to become competitive with state-of-the-art conventional transistors? Then it is not only essential that one can achieve a large wave vector splitting Δk of a spin-split subband but also that it can be done with a small bias. As a benchmark for the performance, we choose the switch energy for Si MOSFETs where 3 aJ has been projected.[44]

We have previously[45] approximated the switch energy for n-type and p-type spin transistors by CV^2, where C is the capacitance of a QW structure surrounded by two gates and V is the applied bias between them. (We have included here turning on and off of the device, which cancels a factor $1/2$). We then concluded that n-type spin transistors with the original design would have problems in becoming competitive with conventional transistors unless fundamentally new ideas were presented.

A similar conclusion was independently drawn by Bandyopadhyay and Cahay.[46] They assumed that a spin transistor must be based on a one-dimensional channel and that only the lowest one-dimensional subband should be filled. However, this resulted in an anomalously small carrier density, 3×10^{10} cm^{-2} or 3×10^{5} cm^{-1}. This made their comparison very unfavorable for the spin transistor.

A more recent comparison with conventional transistors has been made by Hall and Flatté[47] for a modified spin tran-

sistor. It is not based on the Rashba effect but rather on gate-induced spin relaxation. A crucial factor in their approach seems to be efficiency of the gate-induced spin relaxation compared to other spin-relaxation mechanisms. Their comparison was quite favorable for the spin transistor. They estimated a switch energy of 0.5 aJ which is similar to what we find below for our modified spin transistor. The performance of this transistor has been subject to some controversy concerning the need for very efficient spin injection.[48] Since the Hall-Flatté spin transistor is based on another mechanism than the spin precession considered here, it is beyond the scope of the present paper to enter this debate.

Utilizing the built-in electric field in the modulation-doped quantum well, one can achieve a given Δk with a QWB that is an order of magnitude smaller than with a uniform electric field. If we only consider the lowest spin subband pair and follow the approach of Ref. 45, we obtain a switch energy of 0.4 aJ in the modulation-doped case, and 35 aJ in a spin transistor with the same length and uniform electric field. The former figure compares very well with present state-of-the-art transistors. Thus the utilization of the mechanism proposed by us could make a substantial difference for the competitiveness of spin transistors. We have calculated the additional contribution to the switch energy from the redistribution of carriers in the QW, taking the $k_{\|}$-dependent wave functions into account but found that it only increases by about 20%.

A complication with our design is that the second subband pair with the opposite sign of Δk and spin precession direction is also filled. This does not prevent the possibility that the spins at the two interfaces can have made a precession by the angle π but in opposite directions on the arrival to the drain where the transmission becomes low.

It has been demonstrated that one can contact the electron gases in a double quantum well structure separately.[49] It seems feasible that also the interface regions of a wide modulation-doped quantum well can be contacted separately which opens up interesting possibilities occurring from the controllable properties of modulation-doped quantum wells.

One can envision practical problems to create a perfectly symmetric quantum well structure corresponding to the on state of a spin transistor. One possibility is a double-gate structure in which the total carrier concentration and the asymmetry can be controlled separately. In Ref. 50 the back gate voltage was of the order 100 V, which is not very practical for devices. An alternative design[45] is to have a heavily doped semiconductor layer just below the quantum well structure. In this way a larger fraction of the applied voltage falls over the quantum well.

We now turn to the problem of making a spin transistor function, possibly with some modification of the original idea.[7] A fundamental problem is that the Rashba effect can be described in terms of an effective magnetic field that is perpendicular to both the electric field and the direction of motion for the carrier.[1] Even spin-independent scattering leads to a change of the direction of the velocity and thus the axis of the spin precession. It can also be difficult to inject all the carriers in the same direction. In the case we consider in Fig. 3, we obtain a precession length $L=\pi/\Delta k \sim 1$ μm. Ballistic transport over such a distance requires rather low temperatures. An idea with the purpose of balancing the Rashba and Dresselhaus effect[51] by setting $\alpha=\beta$ in the linearized model makes diffusive transport possible but at the price of a substantial transmission in the off state. One-dimensional channels have been proposed in which the carriers are more or less forced to move in the same direction. As mentioned above the small energy separation between the one-dimensional subbands leads to multimode transport for realistic carrier densities. This is not a prohibitive problem as has been demonstrated by Łusakowski *et al.*[52]

A fundamental problem is that in the approach with the Rashba term, which is a reasonable approximation for electrons in an undoped quantum well, the Rashba splitting Δk is proportional to the Rashba coefficient α but the spin decoherence rate becomes proportional to α^2 (Ref. 3). A large α is beneficial for a rapid spin precession and corresponds to a short gate length in a spin transistor but this advantage is thus offset by the shorter spin decoherence time.

An alternate approach has been presented by Bandyopadhyay and Cahay.[53] Instead of relying on the Rashba effect, they propose using the Dresselhaus effect in a structure with a split gate and a parabolic potential. The main reason was to avoid an in-plane magnetic field in the semiconductor from the magnetized source and drain. However, the main requirement for spin precession is that their magnetization is perpendicular to the effective magnetic field in the channel. Thus it can be either along the channel (as drawn in Ref. 53) or perpendicular to the layers (as drawn in Ref. 41). The mechanism in the transistor based on the Dresselhaus effect is changing the bias of the split gate and then it is assumed that the curvature of the parabolic potential changes. Numerical calculations[54] have indicated that, however, the effect of changing the bias is mainly that the potential becomes flatter in the middle when the channel starts to fill while the curvature of the side walls does not change much.

An alternative that has not been given much attention so far is a *p*-type spin transistor. With a suitable design we have shown that one can obtain a large Δk with an electric field as small as 2 kV/cm.[33] The corresponding precession length is only 40 nm and the possibility of having ballistic transport over such a short distance clearly seems feasible. The strong spin-orbit interaction including its dependence on a gate voltage has recently been demonstrated for a *p*-type GaAs/AlGaAs heterostructure by Grbić *et al.*[55] According to our calculations[33] even stronger spin splitting can be achieved for higher hole densities. Furthermore, the strong anisotropy of hole subbands can possibly be utilized to get a preferred direction of motion without lateral confinement. For holes there is no simple relation between spin precession and spin decoherence rates. Estimates based on experimental determinations indicate that the spin decoherence time can be much longer than the transit time. Because of the strongly nonparabolic hole subbands and their mixed heavy-hole and light-hole character, rather cumbersome numerical calculations appear necessary for a more accurate prediction of the transport properties. For small k analytical expressions proportional to k^3 for the Rashba splitting in heavy-hole subbands have been derived.[1] However, it has been found that the largest spin splittings occur beyond the range of validity of this model.[33]

A relevant question is if one can combine the superefficient Rashba effect for holes with the enhancement in

modulation-doped quantum wells presented here. However, we have shown that for *p*-type spin transistors the largest Rashba splitting is obtained for quite small electric fields ($\sim 5-10$ kV/cm) while the effect of modulation doping is to apply a small bias to utilize the built-in electric field of the order 50–100 kV/cm for which the Rashba effect for holes is reduced.

A well-known problem is that the conductivity mismatch between metal and semiconductor severely limits the spin injection efficiency.[56] A proposal by Rashba is having tunnel barriers between the metal contacts and the semiconductor.[57] A fundamental problem recently pointed out by Fert *et al.*[58] is that this decreases the transmission coefficient and increases the dwell time such that it can become long compared to the spin dephasing time in semiconductor-based spin transistors. They instead proposed using carbon nanotubes. It is beyond the scope of the present paper to evaluate the competitiveness of semiconductors vs carbon nanotubes for spintronic devices. However, we would like to point out that this problem occurs for injection from a spin-polarized contact but other solutions in the form of spin filters have been proposed. A particularly interesting idea is to put a magnetic layer on top of a layered semiconductor structure such that the in-plane fringe fields act as a spin filter.[59] The appealing aspect of this solution is that current flows in the channel below the metal without passing any interfaces where the spin polarization can be reduced.

V. DISCUSSION AND CONCLUSIONS

We have implicitly assumed coherence of the wave function across the 80 nm QW with a high and broad barrier in the middle. Whether this coherence actually prevails should depend on the sample quality. This system with our predicted effects seems ideal for further studies of this fundamental problem.

The self-consistent calculations have so far been performed in the Hartree approximation. For studies of spin properties it is conceivable that exchange and correlation can give significant effects, especially in anticrossing situations. This is planned to be examined in future publications.

In our multiband approach the well established Burt-Foreman boundary conditions[12–14] are behind any interface contribution. The exact relation behind this approach and what is obtained by folding down the adjacent bands to the conduction band, as in Refs. 26 and 27, is not trivial and remains to be analyzed. Our approach is in our opinion more sound than one-band approaches that are based on approximations whose accuracy is difficult to determine. Furthermore, the analytical expressions[26,27] are based on consideration of the interfaces, where the EFA has its main weakness and where the actual gradual but rapid potential variation near an interface is replaced by a sharp step. Especially when the interfaces give substantial contributions, it is likely that different operator orderings can influence the results considerably. Calculations with downfolding procedures for wide modulation-doped quantum wells and comparison with the present results would be valuable in evaluating how close the results of these approaches are in a nontrivial case such as this.

It has recently been predicted that the spin Hall effect can be strongly enhanced at a subband anticrossing in a bilayer system.[60] There the potential was not specified but the Rashba coefficients were allowed to differ in the two layers. For further investigations of this effect, modulation-doped quantum wells seem useful due to the possibility of controlling the degree of interaction between the two electron gases and each of the interface fields.

In conclusion we have analyzed the foundations of the envelope function approximation and concluded that, while the Dresselhaus effect should be included as k-dependent terms in a matrix, the proper inclusion of the Rashba effect is adding the macroscopic potential along the diagonal in a multiband approach. This has given good agreement with experiment for two-dimensional hole gases.[21,23,24] The commonly used Rashba term (5) is a hybrid including both potential and k. The proper derivation of such a term within the framework of the EFA with proper boundary conditions[12–14] deserves to be examined more closely.

For symmetric wells with Dresselhaus effect only, we find interesting effects in a modulation-doped quantum well that are qualitatively different from those in a square well.

We have found that, with a nonuniform electric field, insertion of some kind of expectation value or other average into Eq. (6) underestimates the Rashba splitting. Furthermore, this expectation value is not always well-defined for a subband because it can differ substantially between its spin-split components. We have found that the contribution from the interfaces is about half of that from the electric field in the layers for the potential we have considered.

We have demonstrated a very efficient switching mechanism of the Rashba splitting in wide modulation-doped quantum wells. One can use a bias corresponding to a moderate average electric field and still get a Rashba splitting typically enhanced by an order of magnitude due to the built-in local electric field in the interface region. The switching mechanism is based on localization of each wave function to one interface region with a barely sufficient bias. A switching mechanism based on anticrossing in slightly asymmetric quantum wells[42,43] is not included here but will be examined further elsewhere.

The enhancement of the efficiency of the Rashba effect should be valuable for different spintronic devices. We here have focused on spin transistors of the type proposed by Datta and Das.[7] With our modification we find that it can get a potential to outperform conventional transistors. We have also discussed some remaining obstacles to make such spin transistors function.

ACKNOWLEDGMENTS

D.M.G. wishes to acknowledge the hospitality of the School of Information and Communication Technology of the Royal Institute of Technology where some of this work was carried out. Funding has been received from the Serbian Ministry of Science.

*Corresponding author. ulfe@kth.se

[1] R. Winkler, *Spin-Orbit Coupling Effects in Two-Dimensional Electron and Hole Systems*, Springer Tracts in Modern Physics, Vol. 191 (Springer, Berlin, 2003).
[2] Y. A. Bychkov and E. I. Rashba, J. Phys. C **17**, 6039 (1984).
[3] I. Žutić, J. Fabian, and S. Das Sarma, Rev. Mod. Phys. **76**, 323 (2004).
[4] G. Dresselhaus, Phys. Rev. **100**, 580 (1955).
[5] J. M. Luttinger and W. Kohn, Phys. Rev. **97**, 869 (1955).
[6] J. M. Luttinger, Phys. Rev. **102**, 1030 (1956).
[7] S. Datta and B. Das, Appl. Phys. Lett. **56**, 665 (1990).
[8] C. Kittel and A. H. Mitchell, Phys. Rev. **96**, 1488 (1954).
[9] W. Kohn and J. M. Luttinger, Phys. Rev. **98**, 915 (1955).
[10] R. A. Faulkner, Phys. Rev. **184**, 713 (1969).
[11] S. T. Pantelides, Rev. Mod. Phys. **50**, 797 (1978).
[12] M. G. Burt, J. Phys.: Condens. Matter **4**, 6651 (1992), and references therein.
[13] B. A. Foreman, Phys. Rev. B **48**, 4964 (1993).
[14] B. A. Foreman, Phys. Rev. B **56**, R12748 (1997).
[15] A. T. Meney, B. Gonul, and E. P. O'Reilly, Phys. Rev. B **50**, 10893 (1994).
[16] F. Stern and S. Das Sarma, Phys. Rev. B **30**, 840 (1984).
[17] P. O. Löwdin, J. Chem. Phys. **19**, 1396 (1951).
[18] M. Cardona, N. E. Christensen, and G. Fasol, Phys. Rev. Lett. **56**, 2831 (1986).
[19] L. Wissinger, U. Rössler, R. Winkler, B. Jusserand, and D. Richards, Phys. Rev. B **58**, 15375 (1998).
[20] D. Richards, B. Jusserand, G. Allan, C. Priester, and B. Etienne, Solid-State Electron. **40**, 127 (1996).
[21] G. Landwehr, E. Bangert, K. von Klitzing, Th. Englert, and G. Dorda, Solid State Commun. **19**, 1031 (1976).
[22] J. C. Hensel, H. Hasegawa, and M. Nakayama, Phys. Rev. **138**, A225 (1965).
[23] H. L. Stormer, Z. Schlesinger, A. Chang, D. C. Tsui, A. C. Gossard, and W. Wiegmann, Phys. Rev. Lett. **51**, 126 (1983).
[24] U. Ekenberg and M. Altarelli, Phys. Rev. B **32**, 3712 (1985).
[25] E. A. de Andrada e Silva, G. C. LaRocca, and F. Bassani, Phys. Rev. B **55**, 16293 (1997).
[26] W. Zawadzki and P. Pfeffer, Semicond. Sci. Technol. **19**, R1 (2004).
[27] R. Lassnig, Phys. Rev. B **31**, 8076 (1985).
[28] J. Nitta, T. Akazaki, H. Takayanagi, and T. Enoki, Phys. Rev. Lett. **78**, 1335 (1997).
[29] D. Grundler, Phys. Rev. Lett. **84**, 6074 (2000).
[30] B. A. Foreman, Phys. Rev. Lett. **82**, 1339 (1999); O. Krebs and P. Voisin, *ibid.* **82**, 1340 (1999).
[31] O. Krebs and P. Voisin, Phys. Rev. Lett. **77**, 1829 (1996).
[32] W. Yang and K. Chang, Phys. Rev. B **73**, 113303 (2006).
[33] D. M. Gvozdić and U. Ekenberg, Europhys. Lett. **73**, 927 (2006).
[34] D. A. Broido and L. J. Sham, Phys. Rev. B **31**, 888 (1985).
[35] T. Ando, J. Phys. Soc. Jpn. **54**, 1528 (1985).
[36] M. Akabori, V. A. Guzenko, T. Sato, Th. Schäpers, T. Suzuki, and S. Yamada, Phys. Rev. B **77**, 205320 (2008).
[37] D. M. Gvozdić and U. Ekenberg, Appl. Phys. Lett. **90**, 053105 (2007).
[38] R. Winkler and U. Rössler, Phys. Rev. B **48**, 8918 (1993).
[39] U. Ekenberg, Phys. Rev. B **38**, 12664 (1988).
[40] E. Bernardes, J. Schliemann, M. Lee, J. C. Egues, and D. Loss, Phys. Rev. Lett. **99**, 076603 (2007).
[41] R. Winkler, Phys. Rev. B **69**, 045317 (2004).
[42] D. M. Gvozdić and U. Ekenberg, J. Phys.: Conf. Ser. **61**, 273 (2007).
[43] D. M. Gvozdić and U. Ekenberg, Physica E (Amsterdam) **40**, 2081 (2008).
[44] International Technology Roadmap for Semiconductors (ITRS), http://www.itrs.net/Links/2005ITRS/ERD2005.pdf
[45] D. M. Gvozdić, U. Ekenberg, and L. Thylén, J. Supercond. **18**, 349 (2005).
[46] S. Bandyopadhyay and M. Cahay, Appl. Phys. Lett. **85**, 1433 (2004).
[47] K. C. Hall and M. E. Flatté, Appl. Phys. Lett. **88**, 162503 (2006).
[48] S. Bandyopadhyay and M. Cahay, arXiv:cond-mat/0604532 (unpublished); M. E. Flatté and K. C. Hall, arXiv:cond-mat/0607432v1 (unpublished); S. Bandyopadhyay and M. Cahay, arXiv:cond-mat/0607659 (unpublished).
[49] J. M. Eisenstein, L. Pfeiffer, and K. West, Appl. Phys. Lett. **57**, 2324 (1990).
[50] S. J. Papadakis, E. P. De Poortere, H. C. Manoharan, M. Shayegan, and R. Winkler, Science **283**, 2056 (1999).
[51] J. Schliemann, J. C. Egues, and D. Loss, Phys. Rev. Lett. **90**, 146801 (2003).
[52] A. Łusakowski, J. Wróbel, and T. Dietl, Phys. Rev. B **68**, 081201(R) (2003).
[53] S. Bandyopadhyay and M. Cahay, Appl. Phys. Lett. **85**, 1814 (2004).
[54] S. E. Laux, D. J. Frank, and F. Stern, Surf. Sci. **196**, 101 (1988).
[55] B. Grbić, R. Leturcq, T. Ihn, K. Ensslin, D. Reuter, and A. D. Wieck, Phys. Rev. B **77**, 125312 (2008).
[56] G. Schmidt, D. Ferrand, L. W. Molenkamp, A. T. Filip, and B. J. van Wees, Phys. Rev. B **62**, R4790 (2000).
[57] E. I. Rashba, Phys. Rev. B **62**, R16267 (2000).
[58] A. Fert, J.-M. George, H. Jaffrès, and R. Mattana, IEEE Trans. Electron Devices **54**, 921 (2007).
[59] F. Zhai and H. Q. Xu, Appl. Phys. Lett. **88**, 032502 (2006).
[60] P.-Q. Jin and Y.-Q. Li, Phys. Rev. B **76**, 235311 (2007).